

# Long time stability of small amplitude Breathers in a mixed FPU-KG model

Simone Paleari<sup>\*†</sup> and Tiziano Penati<sup>\*‡</sup>

April 27, 2015

## Abstract

In the limit of small couplings in the nearest neighbor interaction, and small total energy, we apply the resonant normal form result of a previous paper of ours to a finite but arbitrarily large mixed Fermi-Pasta-Ulam Klein-Gordon chain, i.e. with both linear and nonlinear terms in both the on-site and interaction potential, with periodic boundary conditions.

An existence and orbital stability result for Breathers of such a normal form, which turns out to be a generalized discrete Nonlinear Schrödinger model with exponentially decaying all neighbor interactions, is first proved.

Exploiting such a result as an intermediate step, a long time stability theorem for the true Breathers of the KG and FPU-KG models, in the anti-continuous limit, is proven.

## 1 Introduction and statement of the results

We consider a mixed Fermi-Pasta-Ulam Klein-Gordon model (FPU-KG) as described by the following Hamiltonian

$$H(x, y) = \frac{1}{2} \sum_{j=1}^N [y_j^2 + x_j^2 + a(x_{j+1} - x_j)^2] + \frac{1}{4} \sum_{j=1}^N [x_j^4 + b(x_{j+1} - x_j)^4] , \quad (1)$$

$$x_0 = x_N , \quad y_0 = y_N , \quad (2)$$

i.e. a finite chain of  $N$  degrees of freedom and periodic boundary conditions, where  $a > 0$  and  $b \geq 0$  are the linear and nonlinear coupling coefficients. It can be remarked that the classical KG model ( $a \neq 0, b = 0$ ) is included as a particular case, but the pure FPU one is clearly not covered<sup>1</sup>.

According to a previous result of ours, for any  $r \geq 1$ , provided the coupling parameters  $a$  and  $b$  are correspondingly small enough, there exists a canonical transformation  $T_{\mathcal{X}}$  which puts the Hamiltonian (1) into an extensive resonant normal form of order  $r$

$$H^{(r)} = H_{\Omega} + \mathcal{Z} + P^{(r+1)} , \quad \{H_{\Omega}, \mathcal{Z}\} = 0 .$$

with  $H_{\Omega}$  a system of  $N$  identical oscillators whose frequency  $\Omega$  turns out to be the average of the linear spectrum of the original Hamiltonian,  $\mathcal{Z}$  a non-homogeneous polynomials of order  $2r + 2$ ,  $P^{(r+1)}$  a remainder of order at least  $2r + 4$  (see Theorem 2.1 in Section 2). Strictly speaking, the original statement is formulated in the case  $b = 0$ , see [17], but holds also for  $b \neq 0$  and small.

---

<sup>\*</sup>Università degli Studi di Milano, Dipartimento di Matematica, Via Saldini 50, 20133 Milano (Italy)

<sup>†</sup>simone.paleari@unimi.it

<sup>‡</sup>tiziano.penati@unimi.it

<sup>1</sup>We can't include the FPU case (which corresponds to  $a \neq 0$  and  $b \neq 0$ , without the on-site potentials  $x_j^2$  and  $x_j^4$ ) since the normal form construction we rely on does not apply for such a model; moreover, in the present context of an application to the stability of Breather solutions, the discussion of such an extension could be pointless since in many FPU-like models, like in [12] no fundamental Breathers exist.

The above normal form was indeed shown in [17] to be well defined in a small ball  $B_R(0)$  both in euclidean and in supremum norm, i.e. both in a regime of finite total and respectively specific<sup>2</sup> energy; one of the key points was indeed to consider finite but arbitrarily large systems (along a direction we followed also, e.g., in [20]), with estimates uniform in the size of the chain, hence valid in the limit  $N \rightarrow +\infty$ . However, only in the case of the euclidean norm the almost invariance of  $H_\Omega$  over times  $|t| \sim R^{-r-1}$  was granted, due to the equivalence between  $H_\Omega$  and the selected norm. Moreover, looking at the structure of  $\mathcal{Z}$ , the normal form  $H_\Omega + \mathcal{Z}$  appears as a generalized discrete nonlinear Schrödinger (GdNLS<sup>3</sup>) chain: it is characterized by all neighbors couplings, with exponential decay of the coefficients with the distance between sites, both in the linear and nonlinear terms, the last ones being of order  $2r + 1 \geq 3$ . Since the Hamiltonian of such a normal form is given by an expansion both in energy, through the degree of the polynomials, and in coupling, it is actually cumbersome and somewhat useless to give here a complete and explicit formulation; the following are the leading terms, in the transformed variables  $(\tilde{x}, \tilde{y})$

$$\begin{aligned}
H_{\text{GdNLS}} = & \sum_{j=1}^N \left[ \frac{\Omega}{2} (\tilde{y}_j^2 + \tilde{x}_j^2) + \right. & H_\Omega \\
& + \mathcal{O}(c) (\tilde{x}_j \tilde{x}_{j+1} + \tilde{y}_j \tilde{y}_{j+1}) + \mathcal{O}(c^2) (\tilde{x}_j \tilde{x}_{j+2} + \tilde{y}_j \tilde{y}_{j+2}) + \mathcal{O}(c^3) + & \mathcal{Z}: \text{quadratic part} \\
& + \mathcal{O}(c^0) (\tilde{x}_j^2 + \tilde{y}_j^2)^2 + \mathcal{O}(c) (\tilde{x}_j^2 + \tilde{y}_j^2) (\tilde{x}_j \tilde{x}_{j\pm 1} + \tilde{y}_j \tilde{y}_{j\pm 1}) + \mathcal{O}(c^2) \Big] + & \mathcal{Z}: \text{quartic part} \\
& + \dots & \mathcal{Z}: \text{higher orders}
\end{aligned}$$

where we have introduced the collective coupling constant

$$c := \max\{a, b\} . \quad (3)$$

If one truncates the above Hamiltonian, using only  $H_\Omega$  and the first term of both the quadratic and quartic part of  $\mathcal{Z}$ , it is possible to recognize the usual dNLS, here written in real coordinates.

In this work we exploit the invariance of  $H_\Omega$  in the above resonant normal form part, in order to get some stability results about true or approximated Breather solution in the model (1). The results we are going to present hold in the small total energy  $E < E_*(r)$  regime and for  $c < c_*(E, r)$  small enough, hence in the anti-continuous limit.

For a more precise formulation, let us give some notation. We denote with  $\mathcal{P}$  the phase space  $\mathbb{R}^{2N}$  endowed by the usual euclidean norm  $\|z\|$ , where  $z = (x, y)$  is the generic element of  $\mathcal{P}$ . Given  $z \in \mathcal{P}$ , let us also denote by  $\mathcal{O}(z)$  the orbit through  $z$ . We also need to introduce a suitable “orbital” distance. We use the Hausdorff distance  $d_H$ , which is a metric once restricted to the subset of non empty and compact sets: since we consider periodic orbits and subsets of orbits parametrized by a closed interval of time,  $d_H$  satisfies all the relevant properties we need. We recall the definition: given two subsets  $A$  and  $B$  of  $\mathcal{P}$ ,

$$\begin{aligned}
d(A, B) &:= \sup_{a \in A} \inf_{b \in B} \|a - b\| , \\
d_H(A, B) &:= \max\{d(A, B), d(B, A)\} .
\end{aligned} \quad (4)$$

Let us denote<sup>4</sup> by  $\Psi_{a,b}$  and  $\mathcal{O}(\Psi_{a,b})$  respectively the Breather initial profile and its orbit for our FPU-KG model (1). The existence of such an object in the anti-continuous limit (i.e. as a

<sup>2</sup>We recall that, given  $E = H(x, y)$  the total energy, the specific energy is the average energy per degree of freedom  $E/N$ .

<sup>3</sup>Recently several works appeared on GdNLS models with more than first neighbor interactions, like [7, 15], or with higher nonlinearities, like [6, 8], where spatially localized periodic orbits, like breathers or multibreathers, are studied. Please remark that we here use the term “generalized” exactly to indicate a *generalization*, without any particular reference to other “generalized” dNLS.

<sup>4</sup>We will follow the following convention to denote profiles, and consequently orbits via the symbol  $\mathcal{O}(\cdot)$ :

$\phi$	generic profile for FPU-KG	$z = (x, y)$	original coordinates
$\Psi$	Breather profile for FPU-KG	$\tilde{z} = (\tilde{x}, \tilde{y})$	transformed coordinates
$\psi$	Breather profile for GdNLS	$\tilde{\psi}$	objects in transformed coordinates

See also (41) for the relation between an object in original coordinates and in transformed ones.

family in  $(a, b)$  emerging from the trivial one-site excitation solution available when  $(a, b) = (0, 0)$  has been obtained originally in [16] (strictly speaking in the case  $b = 0$ ).

In the formulation below of our result, the long stability time  $T_{\epsilon, r, R}$  (see (42)) scales as

$$T_{\epsilon, r, R} \simeq \epsilon^2 \frac{(Rr)^{-2r}}{R^4} , \quad (5)$$

where  $r$  is the aforementioned normal form order,  $R$  control the small energy of the objects involved, and  $\epsilon$ , sufficiently smaller than  $R$ , is the (tunable part of the) radius of the stability neighborhood.

Finally, to unambiguously fix the two-parameter family  $\Psi_{a, b}$ , we require it to emerge, in the anti-continuous limit, from the single-site oscillator with prescribed amplitude  $\|\Psi_0\| = R/6$ , where here and in the following, we will use a single sub-scripted 0 to indicate the values  $(a, b) = (0, 0)$  (see (39) and (53) for explicit definitions).

**Theorem 1.1** *Fix an arbitrary integer  $r \geq 1$ . Then there exists  $R_*(r) < 1$  such that for all  $R < R_*$  and  $0 < \epsilon \ll R^2$  there exist  $c_*(r, R, \epsilon)$  and  $\delta(\epsilon)$ , such that for all  $c < c_*$  the (piece of) orbit  $\mathcal{O}(\phi) := \{\phi(t) : |t| \leq T_{\epsilon, r, R}, \phi(0) = \phi\}$ , solution of (1), satisfies*

$$\|\phi - \Psi_{a, b}\| < \delta \implies d_H(\mathcal{O}(\phi), \mathcal{O}(\Psi_{a, b})) < \epsilon . \quad (6)$$

A first comment on the above statement pertains the coupling threshold  $c_*$  and its dependence on the relevant parameters. It depends on  $r$  because of the normal form construction: the larger the transformation steps number required, the smaller the perturbation parameter, i.e. the coupling. The dependence from  $R$  comes both from the normal form procedure, when we need to control the size of the transformation domains and the smallness of the remainder, and from the existence of Breathers solutions of the GdNLS (this will be shown to be a necessary intermediate step). The  $\epsilon$  dependence appears instead in the last part of the proof, when the distance between the Breather of the full system and that of the normal form must be controlled.

Let us add some more details. As we said, our proof of Theorem 1.1 is based on the long time stability result of Theorem 3.1. We indeed first show the expected existence and stability of a Breather for the normal form (GdNLS), respectively by a continuation from the anti-continuous limit and exploiting the second conserved quantity of the GdNLS. Let us denote by  $\mathcal{O}(\psi_{a, b})$  the orbit of such a Breather, emerging from the same single-site oscillator  $\Psi_0$  introduced above. We remark that the closed trajectory  $\mathcal{O}(\psi_{a, b})$  represents an approximated solution for (1). We then use the small remainder given by the normal form transformation to translate the infinite time stability of the GdNLS dynamics around the GdNLS Breather  $\mathcal{O}(\psi_{a, b})$  into a long time stability of the FPU-KG dynamics around the same object. This concludes the sketch of the proof of Theorem 3.1, where a stability control of the FPU-KG dynamics can be obtained in the form

$$\|\phi - \psi_{a, b}\| < \delta \implies d_H(\mathcal{O}(\phi), \mathcal{O}(\psi_{a, b})) < \epsilon ,$$

for  $|t| \lesssim T_{\epsilon, r, R}$  and  $c < c_*$ , in this case with  $c_*(r, R)$  independent of  $\epsilon$ . Thus, at fixed  $r$  and  $R$ , one can play<sup>5</sup> with  $\epsilon$  to strengthen the stability control without further requirements on the couplings.

To get the result of Theorem 1.1 one eventually exploits the closeness of the FPU-KG Breather  $\mathcal{O}(\Psi_{a, b})$  to the GdNLS Breather  $\mathcal{O}(\psi_{a, b})$ : both the objects emerge in the anti-continuous limit from the same configuration  $\Psi_0$ , thus using the continuity in their (common) continuation parameter  $c$  one gets a (weak) closeness of order  $\sqrt[4]{c}$ . Here enters the dependence of  $c_*$  also on  $\epsilon$ : this is needed in order to ensure that the true Breather configuration  $\Psi_{a, b}$  lies well within the stability basin of the approximated Breather orbit  $\mathcal{O}(\psi_{a, b})$ . Furthermore, in order to transfer the stability of  $\mathcal{O}(\psi_{a, b})$  to the stability of  $\mathcal{O}(\Psi_{a, b})$  the triangular inequality of the Hausdorff metric  $d_H$  is also needed.

We would like to stress that our result resembles, in its formulation and strategy, Theorem 2.1 of [3], which is the first, and actually one of the few, result of long time stability of Breathers

<sup>5</sup>However, in order to get a meaningful result,  $\epsilon$  shouldn't be taken too small: otherwise the stability time  $T_{\epsilon, r, R}$ , which scales as  $\epsilon^2$ , could fall shorter than the period of the (true/approximated) Breather.

for weakly coupled oscillators (see also [4]): indeed, although Nekhoroshev-type stability was expected since the earliest papers (see, e.g. [16]), most of the literature on the stability of Breathers (and of their multi-site generalizations, called Multibreathers) deals with the linear stability (see [2, 18, 19, 27]).

There are however some differences with [3], that we would like to underline here. The first one is that in [3] the closeness to the Breather solutions was obtained with a “local” normal form around a generic an-harmonic oscillator (the system being infinite), using only the linear coupling  $a$  as small parameter (since it treats the model (1) with  $b = 0$ ). As a consequence, it is valid also for arbitrary large amplitudes and not only in the small energy regime, like Theorem 1.1. Our normal form is instead “more global”, in the sense that it holds in a whole neighborhood of the origin. Hence, within its limit of validity given by the smallness of the energy, it can be used to capture the main features of any Cauchy Problem.

Moreover, and differently from our Theorem 1.1, in [3] the small parameter  $a$  is used also in order to fix the domain of stability: indeed, in [3], the corresponding of our radius  $\delta$  of the stability basin vanishes as  $a \rightarrow 0$ . This is a consequence of the way the “local” normal form Theorem (ref. Theorem 4.1 in [3]) has been used, choosing  $\sqrt{a}$  as the size of the domain of validity, and it seems in contrast with the intuition that by approaching the uncoupled system ( $a = 0$  in that case,  $c = 0$  in the present one), the Breather should be increasingly stable, not only in terms of time scale but also in terms of domain. With respect to this aspect, our result is more flexible: as already pointed, at fixed time scale (i.e. fixing  $r$  and  $E$ ) we are allowed to arbitrarily decrease the coupling  $a$  without shrinking the stability basin.

Concerning instead the dependence on the coupling of the stability time scale, the result in [3] appears to be as strong as one could hope, i.e. one has an exponential dependence of the form  $T_a \simeq \exp(a^{-1/6})$ . Our result, on the contrary, seems to fail completely in the expected growth of the time scale as the couplings vanish, since neither  $a$  nor  $b$  appear explicitly in (5). However our result is indeed somewhat similar once the implicit dependence on the couplings is taken into account: the formulation of Theorem 1.1 provides a stability time  $T_{\epsilon, r, R}$  which scales as a power of  $(Rr)^{-1}$ , which is large provided the “amplitude”  $R$  is sufficiently small with respect to  $1/r$  (see also condition (29)), with an exponent  $r$  which can be arbitrarily increased by sufficiently decreasing the coupling  $c$ . In the parameters plane  $(r, c)$ , the allowed region has a border<sup>6</sup> roughly described by  $cr^4 = \text{const.}$  Thus one can either formulate the statement, as we do, assuming an arbitrary  $r$ , provided  $c$  is smaller than something scaling as  $1/r^4$ ; or one could fix  $c$  (sufficiently small for independent reasons) and let  $r$  up to  $1/\sqrt[4]{c}$ . In the latter case, provided  $R$  vanishes at least as  $\sqrt[4]{c}$ , the stability time scale resemble very closely the exponential one of [3]. The price to be paid is that, the smaller is the amplitude  $R$ , the smaller has to be the stability domain parameter  $\epsilon$ .

There is a last comment in the comparison of our results with the reference paper [3]. The stability in [3], as we said, is obtained through a normal form around the an-harmonic oscillator which is going to constitute the core of the Breather, actually by removing the dominant part of the coupling of such an oscillator  $(J, \varphi)$  with the rest of the chain: this typically requires a Diophantine non resonance condition for the true frequency  $\omega(J)$  of the Breather with respect to its linear frequency  $\omega_0$ . However, the smaller is  $J$ , the closer is  $\omega(J)$  to  $\omega_0$  and thus proportionally smaller must be the parameter  $\nu$  in the Diophantine condition  $|k_1\omega + k_2\omega_0| \geq \nu/|k|^2$ . And this affects the small coupling interval  $(0, a_*)$  for which the result in [3] applies: indeed, since the normal form construction needs  $\sqrt{a}/\nu < 1$  in order to be performed, the threshold  $a_*$  has to decrease at least like  $\nu^2$ , which means  $a_* \lesssim R^4$  in terms of small amplitude  $R$ . Our result is instead completely free of any Diophantine condition on the Breather frequency, implicitly requiring only non-resonance and non-degeneracy of the frequency in order to have the existence of the Breather. And we stress that, even though we also require the coupling  $c$  (and then  $a$ ) to be small enough with respect to the amplitude, as a sufficient condition for the variational continuation from the anti-continuous limit, our smallness condition is weaker, being of the order  $c_* \ll R^2$ .

---

<sup>6</sup>We should also remark that different choices for  $\sigma_1$  and  $\sigma_*$ , respectively right before and right after the statement of Theorem 2.1, could lead to a boundary of the form  $cr^\alpha = \text{const.}$ , with  $\alpha > 1$ . This could further improve the time scale dependence on  $c$ , at the price of lowering the thresholds of validity for the relevant parameters (see [17], Section 4.2 for further details); we did not pursue an optimization of this type.

We conclude this Introduction by remarking that, since our strategy is strongly based on a normal form construction for the quadratic part of the Hamiltonian (1) (see also [10,11]), it can be applied also to different local nonlinearities, like for example the Morse or the cubic potential in the DNA models [9,21]. Indeed, even the FPU-KG model presented here is an easy extension with respect to the classical KG one, and we included it here both to give a more general result and because we were motivated by recent papers like [13,27], where a nonlinear quartic interaction is taken into account. Moreover, the perturbation approach we exploited here, even simplified in its preliminary step involving the quadratic part, can be applied to those model where the coupling is purely nonlinear ( $a = 0$ ), thus justifying the long time stability of compact-like Breathers (see [23,24]).

The paper is organized as follows. In Section 2 we reformulate the normal form result (and the main ideas related) discussed in [17]. In Section 3 we present and comment the two results concerning the long time orbital stability of the approximated and true breather solution. A short Appendix contains the proofs of the existence and stability of the GdNLS breather.

## 2 Background: an extensive resonant normal form Theorem

The aim of this Section is to present the resonant normal form Theorem obtained in [17], with a slightly different formulation which is necessary to deduce Theorem 1.1. At variance with respect to the original paper, we here decided to select  $r$ , the order of the normal form, as the main parameter used to express the thresholds of validity of the construction, instead of the small couplings. Such a different choice fixes the order of the normal form, hence its non linear terms, leaving the small couplings as “free” parameters for the continuation procedure from the anti-continuous limit. For the above reasons, and in order to introduce some definitions and remarks necessary for the comprehension of the perturbation part, in this Section we also repeat, and slightly extend, some ingredients of [17].

### 2.1 Extensivity

The perturbation construction developed in [11,17] is strongly based on the property of extensivity typical of a class of Hamiltonian like (1): physically speaking, in all these models the extensivity results from both the translation invariance and the short interaction potentials. In particular, the extensivity allows to sharply manage the dependence on the size of the system  $N$  in the estimates involved in the perturbation approach. We here recall a possible formalization of this property, by means of the cyclic symmetry, which has been already introduced, widely analyzed and then exploited in [10,11,17].

We denote by  $x_j, y_j$  the position and the momentum of a particle, with  $x_{j+N} = x_j$  and  $y_{j+N} = y_j$  for any  $j$ .

**Cyclic symmetry.** We formalize the translation invariance by using the idea of *cyclic symmetry*. The *cyclic permutation* operator  $\tau$ , acting separately on the variables  $x$  and  $y$ , is defined as

$$\tau(x_1, \dots, x_N) = (x_2, \dots, x_N, x_1), \quad \tau(y_1, \dots, y_N) = (y_2, \dots, y_N, y_1). \quad (7)$$

We extend its action on the space of functions as

$$(\tau f)(x, y) = f(\tau(x, y)) = f(\tau x, \tau y).$$

**Definition 2.1** We say that a function  $F$  is cyclically symmetric if  $\tau F = F$ .

We introduce now an operator, indicated by an upper index  $\oplus$ , acting on functions: given a function  $f$ , a new function  $F = f^\oplus$  is constructed as

$$F = f^\oplus := \sum_{l=1}^N \tau^l f. \quad (8)$$

We shall say that  $f^\oplus(x, y)$  is generated by the *seed*  $f(x, y)$ . We shall often use the convention of denoting cyclically symmetric functions with capital letters and their seeds with the corresponding lower case letter.

**Polynomial norms.** Let  $f(x, y) = \sum_{|j|+|k|=s} f_{j,k} x^j y^k$  be a homogeneous polynomial of degree  $s$  in  $x, y$ . Given a positive  $R$ , we define its polynomial norm as

$$\|f\|_R := R^s \sum_{j,k} |f_{j,k}|. \quad (9)$$

**Norm of an extensive function.** Assume now that we are equipped with a norm for our functions  $\|\cdot\|$ , e.g. the above defined polynomial norm. We introduce a corresponding norm  $\|\cdot\|^\oplus$  for an extensive function  $F = f^\oplus$  by defining

$$\|F\|^\oplus = \|f\|, \quad (10)$$

i.e. by actually measuring the norm of the seed. Although the norm so defined depends explicitly on the choice of the seed, this is harmless in the perturbation estimates since

$$\|F\| \leq N \|F\|^\oplus = N \|f\|,$$

for any  $f$  such that  $F = f^\oplus$ .

**Circulant matrices.** Dealing with particular functions which are quadratic forms, the cyclic symmetry coming from extensivity assumes a particular form. Let us thus restrict our attention to the harmonic part of the Hamiltonian: it is a quadratic form represented by a matrix  $A$

$$H_0(x, y) = \frac{1}{2} y \cdot y + \frac{1}{2} A x \cdot x. \quad (11)$$

If the Hamiltonian  $H$  is extensive, then  $H_0 = h_0^\oplus$ . This implies that  $A$  commutes with the matrix  $\tau$  representing the cyclic permutation (7)

$$\tau_{ij} = \begin{cases} 1 & \text{if } i = j + 1 \pmod{N}, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

We remark that the matrix  $\tau$  is orthogonal and generates a cyclic group of order  $N$  with respect to the matrix product.

In our problem the cyclic symmetry of the Hamiltonian implies that the matrix  $A$  of the quadratic form is circulant. Obviously it is also symmetric, so that the space of matrices of interest to us has dimension  $\lfloor \frac{N}{2} \rfloor + 1$ . Indeed, a circulant and symmetric matrix is completely determined by  $\lfloor \frac{N}{2} \rfloor + 1$  elements of its first line.

**Interaction range** We give here a formal characterization of finite and short range interaction, pointing out some properties that will be useful in the rest of the paper. We restrict our analysis to the set of polynomial functions. We start with some definitions. Let us label the variables as  $x_l, y_l$  with  $l \in \mathbb{Z}$ , and consider a monomial  $x^j y^k$  (in multiindex notation).

**Definition 2.2** We define the support  $S(x^j y^k)$  of the monomial and the interaction distance  $\ell(x^j y^k)$  as follows: considering the exponents  $(j, k)$  we set

$$S(x^j y^k) = \{l : j_l \neq 0 \text{ or } k_l \neq 0\}, \quad \ell(x^j y^k) = \text{diam}(S(x^j y^k)). \quad (13)$$

We say that the monomial is left aligned in case  $S(x^j y^k) \subset \{0, \dots, \ell(x^j y^k) - 1\}$ .

The definition above is extended to a homogeneous polynomial  $f$  by saying that  $S(f)$  is the union of the supports of all the monomials in  $f$ , and that  $f$  is left aligned if all its monomials are left aligned. The relevant property is that if  $\tilde{f}$  is a seed of a cyclically symmetric function  $F$ , then there exists also a left aligned seed  $f$  of the same function  $F$ : just left align all the monomials in  $\tilde{f}$ .

**Short range (exponential decay of) interaction.** For the seed  $f$  of a function consider the decomposition

$$f(z) = \sum_{m \geq 0} f^{(m)}(z), \quad f^{(m)}(z) = \sum_{\ell(k) \leq m} f_k z^k, \quad (14)$$

assuming that every  $f^{(m)}$  is left aligned.

**Definition 2.3** *The seed  $f$  (of an extensive function) is said to be of class  $\mathcal{D}(C_f, \sigma)$  if*

$$\|f^{(m)}\|_1 \leq C_f e^{-\sigma m}, \quad C_f > 0, \sigma > 0. \quad (15)$$

**Continuity of extensive polynomials.** We add here some regularity properties, which are absent in [17].

**Lemma 2.1** *Any  $F = f^\oplus$ , polynomial of degree  $m$ , with  $f \in \mathcal{D}(C_f, \sigma)$  is of class  $\mathcal{C}^m(\mathbb{R}^{2N}, \mathbb{R})$ , with*

$$|F(z)| \leq \|F\|^\oplus \|z\|^m. \quad (16)$$

**proof:** Since

$$f(z) = \sum_{|k|=m} f_k z^k,$$

we have also

$$|F(z)| \leq \sum_{j=0}^{N-1} \sum_{|k|=m} |f_k| |z^k \circ \tau^j| \leq \sum_{|k|=m} |f_k| \left( \sum_{j=0}^{N-1} |z^k \circ \tau^j| \right) \leq \|F\|^\oplus \|z\|^m.$$

Any polynomial  $F$  is represented by a symmetric ( $m$ ) multilinear operator  $\hat{F}$ , such that

$$\hat{F}(z, \dots, z) = F(z),$$

hence (16) gives

$$\sup_{\|z\| \leq 1, z \neq 0} |\hat{F}(z, \dots, z)| \leq \|F\|^\oplus < \infty,$$

which is the continuity of  $\hat{F}$ . The continuity of the differentials follows immediately from  $\hat{F}$  being multilinear.  $\square$

**Hamiltonian vector fields** We consider, as an Hamiltonian, an extensive function  $F$  with seed  $f$ ; we will make use of the common notation<sup>7</sup>  $X_F = (X_1, \dots, X_N, X_{N+1}, \dots, X_{2N})$  to indicate the associated Hamiltonian vector field  $J\nabla F$ , with  $J$  given by the Poisson structure. The first easy, but important, result is that also the Hamiltonian vector field inherits, in a particular form, the cyclic symmetry; a possible choice for the equivalent of the seed turn out to be the couple  $(X_1, X_{N+1})$ , i.e. the first and the  $(N+1)^{\text{th}}$  components of the vector. This fact, which will be more clear thanks to the forthcoming Lemma 2.2, allows us to define in a reasonable and consistent way the following norm

$$\|X_F\|_R^\oplus := \|X_1\|_R + \|X_{N+1}\|_R. \quad (17)$$

**Lemma 2.2** *Given  $F = f^\oplus$ , for the components of its Hamiltonian vector field  $X_F$  we have<sup>8</sup>*

$$\begin{aligned} X_j &= \tau^{j-1} X_1 \\ X_{N+j} &= \tau^{j-1} X_{N+1} \end{aligned} \quad j = 1, \dots, N. \quad (18)$$

<sup>7</sup>For an easier notation we drop the Hamiltonian  $F$  in the indexes of the components of the vector field.

<sup>8</sup>An immediate consequence of (18) is that, defining the norm of the vector field as the sum of its components (i.e. a finite  $\ell^1$  norm), we would get  $\|X_F\|_R = N \|X_F\|_R^\oplus$ , which in turn justify the definition (17), and make it consistent with our previous definition (10).

Moreover, it holds

$$\|X_F\|_R^\oplus = \sum_{l=1}^{2N} \left\| \frac{\partial f}{\partial z_l} \right\|_R. \quad (19)$$

## 2.2 Resonant normal form

In this part we recall, with a slightly different statement more based on the parameter  $r$ , the resonant normal form result of [17]. Although the model (1) has an additional nonlinear term, its main Theorem, here formulated as Theorem 2.1 still apply, since it requires some decay properties of the seeds of the Hamiltonian which are true also for (1).

Indeed, we first recall the splitting of the Hamiltonian (1) as a sum of its quadratic and quartic parts  $H = H_0 + H_1$ , i.e.

$$H_0(x, y) := \frac{1}{2} \sum_{j=1}^N [y_j^2 + x_j^2 + a(x_j - x_{j-1})^2], \quad H_1(x, y) := \frac{1}{4} \sum_{j=1}^N [x_j^4 + b(x_{j+1} - x_j)^4]. \quad (20)$$

### 2.2.1 Discussion on the small parameters

Since both  $0 < a < 1$  and  $0 \leq b < 1$  have to be considered as small parameters, we define

$$\mu := \sqrt[4]{\frac{2c}{1+2c}}, \quad (21)$$

where  $c$  has been introduced in (3): the new parameter  $\mu$  will play the role of main perturbation parameter together with the small radius  $R$ . Moreover, in order to deal with exponentially decaying interactions (and to explain why we defined  $\mu$  as we did), let us introduce the following parameters, which will serve as decay rates in the sense of Definition 2.3

$$\sigma_a := \ln \left( \frac{1+2a}{2a} \right), \quad \sigma_b := \ln \left( \frac{1+2b}{2b} \right), \quad \sigma_0 := \min\{\sigma_a, \sigma_b\}. \quad (22)$$

As a consequence of (21) and (22), one has

$$\sigma_0 = \ln \left( \frac{1+2c}{2c} \right), \quad \mu = e^{-\sigma_0/4}. \quad (23)$$

It is important to notice that  $H_1 = h_1^\oplus$  with  $h_1 \in \mathcal{D}(C_{h_1}, \sigma_b)$ . Indeed by definition

$$H_1(x) = h_1^\oplus, \quad h_1 := \pm \frac{1}{4} x_0^4 + \frac{b}{4} (x_1 - x_0)^4;$$

however by rearranging the monomials, it is possible to select  $h_1$  as

$$h_1 = h_1^{(0)} + h_1^{(1)}, \quad h_1^{(0)} = \frac{\pm 1 + 2b}{4} x_0^4, \quad h_1^{(1)} = \frac{3}{2} b x_0^2 x_1^2 - b x_0 x_1 (x_0^2 + x_1^2).$$

with

$$\|h_1^{(l)}\| \leq C_{h_1} e^{-\sigma_b l}, \quad l = 0, 1, \quad C_{h_1} := \frac{7}{4} (1 + 2b),$$

### 2.2.2 Preliminary linear transformation

We start performing the normalization process with an *exponentially localized* linear transformation (see Proposition 2 of [17], and also [10, 11]) to put the quadratic part into a resonant normal form. Rewrite the matrix  $A$ , introduced in (11), as

$$A = (1 + 2a) \left[ \mathbb{I} - \frac{e^{-\sigma_a}}{2} (\tau + \tau^\top) \right], \quad (24)$$



which is clearly circulant and symmetric, and gives a finite range interaction, in the form of a  $e^{-\sigma_a}$  small perturbation of the identity. Let the constant frequency  $\Omega$  be the average of the square roots of the eigenvalues of  $A$ , and take any  $\sigma_1 \in (0, \sigma_0)$ . We have (see mainly Proposition 3.1 in [11] and the related Proposition 1 in [10])

**Proposition 2.1** *The canonical linear transformation  $q = A^{1/4}x$ ,  $p = A^{-1/4}y$  gives the Hamiltonian  $H_0$  the particular resonant normal form*

$$H_0 = H_\Omega + Z_0, \quad \{H_\Omega, Z_0\} = 0 \quad (25)$$

with  $H_\Omega$  and  $Z_0$  cyclically symmetric with seeds

$$h_\Omega = \frac{\Omega}{2}(\tilde{x}_1^2 + \tilde{y}_1^2), \quad \zeta_0 \in \mathcal{D}(C_{\zeta_0}(a), \sigma_0),$$

and transforms  $H_1$  into a cyclically symmetric function with seed

$$h_1 \in \mathcal{D}(C_{h_1}(a), \sigma_1).$$

Some remarks are in order.

1. We first recall that it is exactly the above linear transformation which introduces the interaction among all sites, with an exponential decay with respect to their distance.
2. The original claim in [11] would actually give  $\zeta_0 \in \mathcal{D}(C_{\zeta_0}(a), \sigma_a) \subseteq \mathcal{D}(C_{\zeta_0}(a), \sigma_0)$ , since  $\sigma_a \geq \sigma_0$ . The choice of taking  $\zeta_0 \in \mathcal{D}(C_{\zeta_0}(a), \sigma_0)$ , thus loosing a bit of the exponential decay, is useful to simplify the control of the decay in the whole normal form construction.
3. As in the proof of Lemma 3.4 of [11], the decay rate  $\sigma_1$  of the seed  $h_1$  cannot be equal to that of the linear transformation, but can be chosen arbitrarily close, i.e. one has that  $h_1 \in \mathcal{D}(C_{h_1}(a), \sigma_1)$  for any  $\sigma_1 < \sigma_0$ . This is especially true when  $\sigma_b > \sigma_a = \sigma_0$ . We nevertheless make the following choice for  $\sigma_1$

$$\sigma_1 := \frac{1}{2}\sigma_0, \quad (26)$$

once again, in order to simplify some calculations.

### 2.2.3 Normal form Theorem

From now on we will simply indicate with  $C$  any constant which does not depend on the relevant parameters, i.e.  $R$ ,  $r$  and  $c$ . Consider the extensive Hamiltonian  $H$  in the new “exponentially localized” coordinates  $(\tilde{x}, \tilde{y})$ , introduced by the previous linear transformation

$$H = H_\Omega + Z_0 + H_1;$$

we have (see Theorem 1 in [17]):

**Theorem 2.1** *Consider the Hamiltonian  $H = h_\Omega^\oplus + \zeta_0^\oplus + h_1^\oplus$  with seeds  $h_\Omega = \frac{\Omega}{2}(x_0^2 + y_0^2)$ , the quadratic term  $\zeta_0$  of class  $\mathcal{D}(C_{\zeta_0}, \sigma_0)$  with  $\zeta_0^{(0)} = 0$ , and the quartic term  $h_1$  of class  $\mathcal{D}(C_{h_1}, \sigma_1)$ . Pick a positive  $\sigma_0/4 \leq \sigma_* < \sigma_1$ , then there exist positive  $\gamma$ ,  $r_*$  and  $C_*$  such that for any positive integer  $r$  satisfying*

$$r < r_*, \quad (27)$$

*there exists a finite generating sequence  $\mathcal{X} = \{\chi_1^\oplus, \dots, \chi_r^\oplus\}$  of a Lie transform such that  $T_{\mathcal{X}}H^{(r)} = H$  where  $H^{(r)}$  is an extensive function of the form*

$$H^{(r)} = H_\Omega + \mathcal{Z} + P^{(r+1)}, \quad \begin{aligned} \mathcal{Z} &:= Z_0 + \dots + Z_r \\ L_\Omega Z_s &= 0, \quad \forall s \in \{0, \dots, r\}, \end{aligned} \quad (28)$$

with  $Z_s$  of degree  $2s+2$  and  $P^{(r+1)}$  a remainder starting with terms of degree equal or bigger than  $2r+4$ .

Moreover, defining  $C_r := 64r^2 C_*$  and  $\sigma_j := \sigma_1 - \frac{j-1}{r}(\sigma_1 - \sigma_*)$ , the following statements hold true:

- (i) the seed  $\chi_s$  of  $\mathcal{X}_s$  is of class  $\mathcal{D}(C_r^{s-1} \frac{C_{h_1}}{\gamma_s}, \sigma_s)$ ;
- (ii) the seed  $\zeta_s$  of  $Z_s$  is of class  $\mathcal{D}(C_r^{s-1} \frac{C_{h_1}}{s}, \sigma_s)$ ;
- (iii) with the choice  $\sigma_* = \sigma_0/4$ , if it is satisfied the smallness condition on the total energy

$$R < R^* := \sqrt{\frac{2}{3(1+e)C_r}}, \quad (29)$$

then the generating sequence  $\mathcal{X}$  defines an analytic canonical transformation on the domain  $B_{\frac{2}{3}R}$  with the properties

$$B_{R/3} \subset T_{\mathcal{X}} B_{\frac{2}{3}R} \subset B_R \quad B_{R/3} \subset T_{\mathcal{X}}^{-1} B_{\frac{2}{3}R} \subset B_R .$$

Moreover, the deformation of the domain  $B_{\frac{2}{3}R}$  is controlled by

$$z \in B_{\frac{2}{3}R} \quad \Rightarrow \quad \|T_{\mathcal{X}}(z) - z\| \leq CC_* R^3, \quad \|T_{\mathcal{X}}^{-1}(z) - z\| \leq CC_* R^3. \quad (30)$$

- (iv) with the choice  $\sigma_* = \sigma_0/4$ , if it is satisfied (29), the remainder is an analytic function on  $B_{\frac{2}{3}R}$ , and it is represented by a series of extensive homogeneous polynomials  $H_s^{(r)}$  of degree  $2s+2$

$$P^{(r+1)} = \sum_{s \geq r+1} H_s^{(r)} \quad H_s^{(r)} = \left(h_s^{(r)}\right)^{\oplus}, \quad (31)$$

and the seeds  $h_s^{(r)}$  are of class  $\mathcal{D}(2\tilde{C}_r^{s-1} C_{h_1}, \sigma_*)$  with  $\tilde{C}_r = \frac{3}{2}C_r$ .

In the following, for the same reasons bringing to the choice (26), we will assume  $\sigma_* = \sigma_0/4$ , as in the last two sub-points of Theorem 2.1. Hence, from (23) and the previous setting of  $\sigma_*$  one gets the relation

$$\mu = e^{-\sigma_*}, \quad (32)$$

and it is possible to give the following values of some of the constants involved in the above Theorem:

$$\begin{aligned} r_* &= \frac{\Omega}{24C_{\zeta_0}} f(\mu), \quad f(\mu) := \frac{(1-\mu^4)(1-\mu^3)}{\mu^2} \\ \gamma &= 2\Omega \left(1 - \frac{r}{2r_*}\right), \\ C_* &= \frac{3C_{h_1}}{\gamma(1-\mu^4)(1-\mu^3)}. \end{aligned} \quad (33)$$

By noticing that condition (27) implies

$$\Omega < \gamma < 2\Omega,$$

we obtain that  $C_*$  essentially depends on  $\mu$ , through  $\sigma_0$  and  $\frac{C_{h_1}}{\Omega}$

$$\frac{3C_{h_1}}{2\Omega(1-\mu^4)(1-\mu^3)} < C_* < \frac{3C_{h_1}}{\Omega(1-\mu^4)(1-\mu^3)};$$

and this provides  $C_r = C_r(r, \mu)$  and  $R^* = R^*(r, \mu)$  with

$$\frac{\partial C_r}{\partial r} > 0, \quad \frac{\partial C_r}{\partial \mu} > 0, \quad \frac{\partial R^*}{\partial r} < 0, \quad \frac{\partial R^*}{\partial \mu} < 0.$$

In the forthcoming application, developed in Section 3, instead of fixing  $\mu$  as the main (small) parameter like in Theorem 2.1, we decide to pick the order  $r \geq 1$  of the normal form, and thus the length of the time scale, as the principal parameter.

As a consequence, by inverting the function  $f(\mu)$  in the first of (33), the normal form (28) holds for all  $\mu < \mu^*(r)$ , with

$$\mu^*(r) := f^{-1}\left(\frac{24C_{\zeta_0}r}{\Omega}\right). \quad (34)$$

Thus, for any  $\mu < \mu^*$  we have  $R^*(r, \mu) > R^*(r, \mu^*)$ . We then take a threshold  $R_*(r)$  for the norm which is uniform with  $\mu < \mu^*$

$$R_*(r) := R^*(r, \mu^*(r)). \quad (35)$$

We summarize the new conditions on the parameters as follows

$$\begin{aligned} r &\geq 1, \\ \mu &< \mu_1^*(r), \quad \Leftrightarrow \quad c < c_1^*(r), \\ R &< R_*(r). \end{aligned} \quad (36)$$

The normal form Theorem 2.1 immediately gives the almost invariance of  $H_\Omega$  and  $\mathcal{Z}$ , which we here formulate (see [17], proof of Corollary 1) in the transformed variables  $\tilde{z} = \mathcal{T}_\mathcal{X}(z)$

**Corollary 2.1** *Let us take  $\tilde{z}(0) \in B_{\frac{4}{3}R}$  and let  $\tau > 0$  be the escape time of the orbit  $\tilde{z}(t)$  from  $B_{\frac{2}{3}R}$ . Then, for all times  $|t| < \tau$ , the approximate integrals of motion  $H_\Omega$  and  $\mathcal{Z}$  fulfill*

$$\begin{aligned} |H_\Omega(\tilde{z}(t)) - H_\Omega(\tilde{z}(0))| &\leq C \frac{C_{h_1}\Omega}{(1-\mu)^2} R^4 \left(\frac{2}{3}R^2C_r\right)^r |t|, \\ |\mathcal{Z}(\tilde{z}(t)) - \mathcal{Z}(\tilde{z}(0))| &\leq C \frac{C_{h_1}(\mu C_{\zeta_0} + C_{h_1}R^2)}{(1-\mu)^2} R^4 \left(\frac{2}{3}R^2C_r\right)^r |t|. \end{aligned}$$

### 3 Stability of true and approximated FPU-KG breathers

Let us now denote the normal form part of  $H^{(r)}$  – see (28) – as

$$K := H_\Omega + \mathcal{Z}, \quad \{H_\Omega, \mathcal{Z}\} = 0, \quad (37)$$

so that the transformed Hamiltonian  $H^{(r)}$  can be split as  $H^{(r)} = K + P^{(r+1)}$ , and the corresponding Hamilton equations are

$$\dot{z} = X_K(z) + X_{P^{(r+1)}}(z). \quad (38)$$

The Hamiltonian  $K$  (the normal form) looks naturally as the Hamiltonian of a Generalized discrete Non Linear Schrödinger equation (GdNLS), with  $H_\Omega$  in the role of the additional conserved quantity; an explicit expression of the leading terms of  $K$  is available in the Introduction.

As a first, and intermediate, application of such a normal form, we give an approximation result for the original system (1): we show that for sufficiently small couplings its dynamics stays close for long times to a closed trajectory in the phase space, provided its initial datum is also close enough to such an object. This trajectory is not an orbit of the original system, but it is a breather of the GdNLS model. The theorem we formulate actually contains, as a first point, and then exploits, an existence and stability result for the GdNLS breather itself with respect to the GdNLS dynamics. Such a first part, despite the generalized nature of the model, is not unexpected in the anti-continuous limit. The other point, actually the long time control for the (FPU-)KG model, is less trivial and indeed it is strongly based on our normal form result.

As a second application, we obtain a result of stability of the true breather for the original system (1), based on the observation that in the anti-continuous limit there always exists a true breather which is close enough, with respect to the greatest parameter  $c$ , to the approximated one. Thus, the stability we get is actually due to the stability of the GdNLS orbit.

### 3.1 Stability of approximated FPU-KG breathers

Since we base the existence part on the anti-continuous limit, let us denote by  $\tilde{\psi}_0$  the 0<sup>th</sup>-site excitation in the transformed coordinates  $(\tilde{x}_j, \tilde{y}_j)$ , i.e.

$$\tilde{\psi}_0 := \{(\tilde{x}_j, \tilde{y}_j)_{j=0, \dots, N-1} : \tilde{x}_0 = \rho, \tilde{y}_0 = 0, \tilde{x}_j = \tilde{y}_j = 0 \ \forall j \neq 0\}, \quad (39)$$

which is indeed the profile of an initial datum belonging to a periodic orbit  $\mathcal{O}(\tilde{\psi}_0)$  (trivially a breather) for the uncoupled system with  $a = b = \mu = 0$  (see (47) below), and for every fixed value of  $\rho$ . A consistent choice for the values of  $\rho$  will be made later.

**Theorem 3.1** *Given  $r$  and  $R$  fulfilling (36), there exists  $c_*(r, R)$  such that, for any  $c < c_*$ :*

1. *there exist a profile  $\tilde{\psi}_{a,b}$  and a frequency  $\lambda_{a,b}$  such that  $\tilde{\psi}_{a,b}e^{i\lambda_{a,b}t}$  is a Breather solution for the GdNLS (37) with  $\|\tilde{\psi}_{a,b}\| = R/6$  and*

$$\|\tilde{\psi}_{a,b} - \tilde{\psi}_0\| \leq C\mu. \quad (40)$$

2. *let us define*

$$\psi_{a,b} := \mathcal{T}_{\mathcal{X}}^{-1} \tilde{\psi}_{a,b}. \quad (41)$$

*For any  $0 < \epsilon \ll R^2$  there exists  $\delta(\epsilon)$  such that the (piece of) orbit  $\mathcal{O}(\phi) := \{\phi(t) : |t| \leq T_{\epsilon, r, R}, \phi(0) = \phi\}$ , solution of (38), satisfies*

$$\|\phi - \psi_{a,b}\| < \delta \implies d_H(\mathcal{O}(\phi), \mathcal{O}(\psi_{a,b})) < \epsilon.$$

*where*

$$T_{\epsilon, r, R} := C_T \frac{\epsilon^2}{R^4} (C_{**} R r)^{-2r}, \quad (42)$$

*with  $C_T$  a suitable constant independent on  $\epsilon$ ,  $r$  and  $R$ , and  $C_{**} := 8\sqrt{\frac{2C_*}{3}}$ .*

We could rephrase the result as follows: for small but non vanishing coupling  $\mu$ , if we start close enough to the trajectory  $\mathcal{O}(\psi_{a,b})$  of a GdNLS Breather, we stay close to it (actually in a small tubular neighborhood of it) for long times.

The proof of the Theorem is made of three steps, which are discussed in the following subsections: first the existence of a breather for the GdNLS with a continuation from the  $\mu = 0$  limit, then its orbital stability exploiting the exact conservation of  $H_\Omega$  (or equivalently of  $\mathcal{Z}$ ) for the Hamiltonian  $K$ , and as a last step the control of the time scale needed to see the effect of the remainder  $P^{(r+1)}$  once the dynamics taken into account is that of the original system.

#### 3.1.1 Existence of Breather solutions for the GdNLS

We denote with  $\mathcal{S} \subset \mathcal{P}$  the sphere  $\mathcal{S} := \{z \in \mathcal{P} \mid H_\Omega(z) = \rho^2\}$  of (small) radius  $\rho < R_*$ . The proof of the first part of Theorem 3.1 is given by the Proposition below, setting  $\rho = R/6$ .

**Proposition 3.1** *Given  $\rho < R_*$ , there exists a threshold  $c_2^*(\rho)$  and a function  $G : (a, b) \mapsto \tilde{\psi}_{a,b}$ , which belongs to  $\mathcal{C}^1([0, c_2^*] \times [0, c_2^*], \mathcal{S})$ , such that  $G(0, 0) = \tilde{\psi}_0$  and*

$$d\mathcal{Z}|_{\mathcal{S}}(\tilde{\psi}_{a,b}, a, b) = 0. \quad (43)$$

*Moreover,  $\tilde{\psi}_{a,b}$  is close to  $\tilde{\psi}_0$*

$$\|\tilde{\psi}_{a,b} - \tilde{\psi}_0\| \leq C\mu. \quad (44)$$

A formal proof of Proposition 3.1 is deferred to the Appendix. As we said above, the idea for existence and localization is to exploit a continuation from the uncoupled limit. If  $\mu = 0$ , the model  $K$  reduces to a system of  $N$  uncoupled an-harmonic oscillators which admits  $\tilde{\psi}_0$  (see (39) above) as a local extremizer of the constrained problem

$$\lambda X_{H_\Omega}(z) = X_{\mathcal{Z}}(z) , \quad \mathcal{Z} := Z_1 + \dots + Z_r , \quad (45)$$

where

$$\zeta_s(\tilde{x}, \tilde{y}) := c_s(\tilde{x}_0^2 + \tilde{y}_0^2)^{s+1} .$$

Indeed, for  $\mu = 0$  (which means  $a = b = 0$ ), the first linear transformation becomes the identity and all the resonant normal form construction reduces to  $N$  identical Birkhoff normal forms for a single an-harmonic oscillator. This means that  $\sigma_0 = \sigma_* = \infty$  and  $\gamma = \Omega = 1$  and  $c_s$  fulfill

$$\begin{cases} |c_1| = C_{h_1} \\ |c_s| \leq \frac{C_{h_1}}{s} C_r^{s-1} , \quad s = 2, \dots, r \end{cases} . \quad (46)$$

Moreover, from its definition in Theorem 2.1,  $C_r = \mathcal{O}(r^2 C_{h_1})$ . The uncoupled constrained problem (45) reads explicitly

$$\begin{cases} 2\lambda_0 \tilde{x}_j = 4\tilde{x}_j \sum_{s=1}^r s c_s (\tilde{x}_j^2 + \tilde{y}_j^2)^s \\ 2\lambda_0 \tilde{y}_j = 4\tilde{y}_j \sum_{s=1}^r s c_s (\tilde{x}_j^2 + \tilde{y}_j^2)^s \\ \sum_j \tilde{x}_j^2 + \tilde{y}_j^2 = \rho^2 \end{cases} , \quad \Rightarrow \quad \tilde{\psi}_0(t) = \tilde{\psi}_0 e^{i\lambda_0 t} , \quad (47)$$

with the Lagrange multiplier  $\lambda_0 = 2 \sum_{s=1}^r s c_s \rho^{2s}$  satisfying

$$|\lambda_0| \geq 2C_{h_1} \rho^2 - 2 \sum_{s=2}^r s |c_s| \rho^{2s} > \frac{5}{7} C_{h_1} \rho^2 , \quad (48)$$

where (46) and the condition (29) have been used. The constrained Hessian  $M$  is a block-diagonal matrix with blocks  $M_j$ ; when evaluated on  $\tilde{\psi}_0$ , for any  $j \neq 0$  the block is  $M_j(\tilde{\psi}_0) = -2\lambda_0 \mathbb{I}$ , while for  $j = 0$  its spectrum is  $\sigma(M_0(\tilde{\psi}_0)) = \{0, \mathcal{O}(C_{h_1} \rho^2)\} = \{0, 4\lambda_0\}$  with associated eigenvectors  $(-\tilde{y}_0, \tilde{x}_0)$  for the Kernel and  $(\tilde{x}_0, \tilde{y}_0)$  for the positive direction. This easily proves that  $M$  is definite in all the directions transverse to the orbit generated by the Hamiltonian field  $X_{H_\Omega} = \Omega(-\tilde{y}_0, \tilde{x}_0, 0, \dots, 0)$ . Then<sup>9</sup> for  $\mu/\rho^2$  small enough the Implicit Function Theorem (IFT) can be applied to uniquely continue  $\tilde{\psi}_0$  to a solution  $\tilde{\psi}_{a,b}$  of  $\mathcal{Z}$  constrained to the level surfaces of  $H_\Omega$  which generates a breather evolution  $\mathcal{O}(\tilde{\psi}_{a,b})$

$$\begin{cases} \lambda_{a,b} \nabla H_\Omega = \nabla \mathcal{Z} \\ \sum_j \tilde{x}_j^2 + \tilde{y}_j^2 = \rho^2 \end{cases} , \quad \Rightarrow \quad \mathcal{O}(\tilde{\psi}_{a,b}) = \tilde{\psi}_{a,b} e^{i\lambda_{a,b} t} . \quad (49)$$

Concerning the exponential localization, we first remark that it is a meaningful property also in finite dimension since we aim at estimates uniform in  $N$ . From a technical point of view, in the infinite dimensional case ( $N = \infty$ ) the exponential decay of the amplitudes  $\tilde{x}_j^2 + \tilde{y}_j^2$  of (49) can be obtained for example with the IFT on the (Hilbert) phase space  $\ell_\sigma^2 \times \ell_\sigma^2$  of square summable sequences with an exponential decay. Alternatively, it can be obtained by homoclinic orbits, as in [22], or by some properties of the inverse of a Tridiagonal linear operator (still in the IFT framework), as in [16].

In the finite case ( $N < \infty$ ), we can speak of an asymptotic exponential decay as  $N$  grows arbitrarily large. For the sake of simplicity, we prefer showing the proof in the energy norm. However, we stress here that the proof holds also when using a norm with exponential weights, i.e. the finite dimensional subspace of  $\ell_\sigma^2$ . In this case, the coercivity constant of the constrained extremizer is proportional to  $e^{-2\sigma}$ , hence the threshold  $\mu^*(\rho, \sigma)$  will be a decreasing function of the decay rate  $\sigma$ .

<sup>9</sup>As clearly explained for example in [26], Section 5.

### 3.1.2 Orbital stability of the GdNLS breather

**Proposition 3.2** *Let  $\tilde{\psi}_{a,b}$  be given by Proposition 3.1. For any positive  $\epsilon \ll \rho^2$  there exists a positive  $\delta(\epsilon)$  such that the orbit  $\mathcal{O}(\tilde{\phi}) := \{\tilde{\phi}(t) : t \in \mathbb{R}, \tilde{\phi}(0) = \tilde{\phi}\}$  of the GdNLS (see (37)), satisfies*

$$\left\| \tilde{\phi} - \tilde{\psi}_{a,b} \right\| < \delta \implies d_H(\mathcal{O}(\tilde{\phi}), \mathcal{O}(\tilde{\psi}_{a,b})) < \epsilon .$$

Given the above use of a constrained critical point formulation to get the existence of the Breather, the Lyapunov stability in the energy norm of such an orbit follows once we verify that the Breather is still an extremizer of  $\mathcal{Z}$  in all the constrained transverse directions and we exploit the fact that  $H_\Omega$  and  $\mathcal{Z}$  are exact constants of motion for  $K$ . The detailed proof is deferred to the Appendix.

### 3.1.3 Orbital stability of the approximated FPU-KG breather

Here<sup>10</sup> we want to use the existence and stability of  $\tilde{\psi}_{a,b}$  of Propositions 3.1 and 3.2 in order to prove the second part of Theorem 3.1. For this reason, we start taking

$$c_*(r, R) := \min\{c_1^*, c_2^*\} , \quad (50)$$

where  $c_{1,2}^*$  are introduced in (36) and in Proposition 3.1.

We initially remark that the result is first formulated for the transformed Hamiltonian (37). In order to give the corresponding statement in the original variables, one has to recall that the canonical transformation  $\mathcal{T}_\mathcal{X}$  is a perturbation of the identity (30), hence

$$\mathcal{T}_\mathcal{X}(z) = z + w(z) \quad \|w(z)\| = \mathcal{O}(\|z\|^3) ,$$

which implies that  $\mathcal{T}_\mathcal{X}$  (and its inverse) is locally Lipschitz in any ball  $B_R$  sufficiently close to the origin, with a constant  $L = \mathcal{O}(1)$ . Thus the control in the transformed variables can be transferred to the original ones using the Lipschitz constant of  $\mathcal{T}_\mathcal{X}^{-1}$

$$d(\phi(t), \mathcal{O}(\psi_{a,b})) = \inf_{\tilde{w} \in \mathcal{O}(\tilde{\psi}_{a,b})} \left\| \mathcal{T}_\mathcal{X}^{-1}\tilde{\phi}(t) - \mathcal{T}_\mathcal{X}^{-1}\tilde{w} \right\| \leq L d(\tilde{\phi}(t), \mathcal{O}(\tilde{\psi}_{a,b})) .$$

Let us work, then, in the transformed variables  $\tilde{z}$ . Our original system is in the normal form (38), thus  $H_\Omega$  and  $\mathcal{Z}$  are only approximate integrals of motion. Hence the drift from the tubular neighborhood of  $\mathcal{O}(\tilde{\psi}_{a,b})$  is bounded by the variation of  $H_\Omega$  and  $\mathcal{Z}$ . We assume  $\phi(0) \in B_{R/3}$  and  $\|\tilde{\psi}_{a,b}\| = R/6$  (hence  $\tilde{\psi}_{a,b} \in B_{\frac{2}{3}R}$ ) where  $R$  satisfies (36). We define

$$\tilde{\phi}(t) := \mathcal{T}_\mathcal{X}\phi(t) , \quad \tilde{\phi}(0) := \mathcal{T}_\mathcal{X}\phi(0) .$$

The variations of  $H_\Omega$  and  $\mathcal{Z}$  in the transformed variables are controlled by Corollary 2.1, at least as long as  $\tilde{\phi}(t) \in B_{\frac{2}{3}R}$ .

The idea is to work as in the proof of Proposition (3.2) in the Appendix. Let us assume that  $\tilde{\phi}(t) \in \mathcal{U}$  (which is true for  $t < T_\mathcal{U}$  for some  $T_\mathcal{U}$  if  $\tilde{\phi}(0)$  is sufficiently close to  $\tilde{\psi}_{a,b}$ ), where  $\mathcal{U}$  is the tubular neighborhood (defined in the above mentioned proof). In such a neighborhood the orbital distance from  $\tilde{\psi}_{a,b}$  can be related to the variations of  $H_\Omega$  and  $\mathcal{Z}$  as in (67). Then it follows immediately

$$\begin{aligned} d(\tilde{\phi}(t), \mathcal{O}(\tilde{\psi}_{a,b})) &< \sqrt{c_3 |H_\Omega(\tilde{\phi}(t)) - H_\Omega(\tilde{\phi}(0))| + \frac{c_4}{C_\mu} |\mathcal{Z}(\tilde{\phi}(t)) - \mathcal{Z}(\tilde{\phi}(0))|} + \\ &+ \sqrt{c_3 |H_\Omega(\tilde{\phi}(0)) - H_\Omega(\tilde{\psi}_{a,b})| + \frac{c_4}{C_\mu} |\mathcal{Z}(\tilde{\phi}(0)) - \mathcal{Z}(\tilde{\psi}_{a,b})|} =: A + B . \end{aligned}$$

<sup>10</sup>From this subsection on, we will often use  $d(A, B)$ , as defined in (4), in the particular case of the set  $A$  given by a single point.

with  $C_\mu$  defined at the beginning of the proof of Proposition 3.2. By using both  $(C_{\zeta_0}\mu + C_{h_1}R^2)/C_\mu = \mathcal{O}(1)$  and Corollary 2.1, we obtain

$$A^2 \leq C \frac{C_{h_1}\Omega}{(1-\mu)^2} R^4 \left( \frac{2}{3} R^2 C_r \right)^r |t| , \quad (51)$$

which gives

$$A < \frac{\epsilon}{2L} , \quad |t| \leq T_{\epsilon,r,R} ,$$

with a suitable choice of  $C_T$ . On the other hand, the distance in the original coordinates can be bounded by exploiting the (local) Lipschitz constant  $L$

$$d(\tilde{\phi}(0), \tilde{\psi}_{a,b}) \leq Ld(\phi(0), \psi_{a,b}) \leq L\delta(\epsilon) ,$$

thus for any  $\epsilon$  there exists  $\delta(\epsilon)$  such that

$$d(\phi(0), \psi_{a,b}) < \delta \quad \Rightarrow \quad B < \frac{\epsilon}{2L} .$$

We can collect all the previous estimates to get

$$d(\tilde{\phi}(t), \mathcal{O}(\tilde{\psi}_{a,b})) < \frac{\epsilon}{L} ,$$

which ensures  $\tilde{\phi}(t) \in \mathcal{U}$  and yields to

$$d(\phi(t), \mathcal{O}(\psi_{a,b})) < \epsilon \ll R^2 , \quad (52)$$

for all  $|t| \leq T_{\epsilon,R,r}$ , i.e. we have  $d(\mathcal{O}(\phi), \mathcal{O}(\psi_{a,b})) < \epsilon$ . The same arguments of the final part of the proof of Proposition 3.2 apply here, so we can get also  $d(\mathcal{O}(\psi_{a,b}), \mathcal{O}(\phi)) < \epsilon$  and conclude the estimate.

As a final comment, we observe that for the time scale considered, the orbits  $\mathcal{O}(\phi)$  and  $\mathcal{O}(\tilde{\phi})$  remain in the domains of definition of the normal form. Indeed since

$$\left\| \mathcal{T}_{\mathcal{X}}^{-1} \mathcal{O}(\tilde{\psi}_{a,b}) \right\| = \left\| \tilde{\psi}_{a,b} \right\| + \mathcal{O}(R^3) = \frac{R}{6} + \mathcal{O}(R^3) ,$$

we obtain

$$\|\phi(t)\| < \frac{R}{3} , \quad \|\tilde{\phi}(t)\| < \frac{2}{3}R .$$

□

### 3.2 Orbital stability of the true FPU-KG Breather: proof of Theorem 1.1

We recall again that when  $c = 0$  the normal form transformation  $\mathcal{T}_{\mathcal{X}}$  corresponds to the common Birkhoff change of coordinates replicated for all the identical an-harmonic oscillators. We denote by

$$\Psi_0 := \mathcal{T}_{\mathcal{X}}^{-1} \tilde{\psi}_0 , \quad (53)$$

the one-site excitation in the Birkhoff coordinates, with  $\tilde{\psi}_0$  given by (39). Since

$$\left\| \Psi_0 - \tilde{\psi}_0 \right\| \approx \left\| \tilde{\psi}_0 \right\|^3 ,$$

the new amplitude will be a  $R^3$  deformation of the original one

$$\|\Psi_0\| = \frac{R}{6} + \mathcal{O}(R^3) .$$

At fixed small amplitude, when the coupling parameters  $a, b$  are switched on, the one-site periodic orbit  $\Psi_0$  can be continued to (form) a family  $\Psi_{a,b}$ , provided  $c < c_3^*(R)$ , as originally proved in [16]. On the other hand, the reference solution  $\tilde{\psi}_0$  can be continued to (form) a family  $\tilde{\psi}_{a,b}$  of orbitally stable Breather solutions for the normal form (37), as claimed in Proposition 3.2. We recall that in (41) we have denoted by  $\psi_{a,b}$  the inverse image of such a family of *approximated Breather solution* for the original FPU-KG model, in the sense of Proposition 3.1. Due to our initial choice for  $\Psi_0$ , the two families initially coincide. Hence, there exists  $c_4^*(r, R) < c_3$ , such that for  $c < c_4^*$  the two families are  $\mu$ -close

$$\|\Psi_{a,b} - \psi_{a,b}\| < C\mu. \quad (54)$$

With this kind of control we are actually able to close the proof by the use of the triangle inequality (and this is ultimately the reason to use the Hausdorff distance). Indeed we control the distance between  $O(\Psi_{a,b})$  and  $O(\phi)$  triangulating via  $O(\psi_{a,b})$ .

We proceed as follow. We exploit the stability of the GdNLS breathers to control both  $O(\phi)$  and  $O(\Psi_{a,b})$ . We thus apply twice the second part of Theorem 3.1, first to  $O(\phi)$ :

$$\exists \delta(\epsilon/2) : \|\phi - \psi_{a,b}\| < \delta \implies d_H(O(\psi_{a,b}), O(\phi)) < \frac{\epsilon}{2}, \quad (55)$$

and then to  $O(\Psi_{a,b})$ :

$$\exists \delta(\epsilon/2) : \|\Psi_{a,b} - \psi_{a,b}\| < \delta \implies d_H(O(\psi_{a,b}), O(\Psi_{a,b})) < \frac{\epsilon}{2}. \quad (56)$$

Concerning this second estimate we remark that the period of the Breather  $O(\Psi_{a,b})$  is of order 1 and surely shorter<sup>11</sup> than then stability time  $T_{\epsilon,r,R}$ ; the Breather itself is thus entirely contained in the tubular neighborhood.

In order to use implications (55) and (56) one has to ensure the control on the distance of the initial datum from  $\psi_{a,b}$ : for  $\Psi_{a,b}$  we use (54), and for  $\phi$  we triangulate again, this time around  $\Psi_{a,b}$ . More precisely

$$\|\phi - \psi_{a,b}\| \leq \|\phi - \Psi_{a,b}\| + \|\Psi_{a,b} - \psi_{a,b}\| < \delta(\epsilon/2),$$

where the first addendum is the one whose smallness we are free to impose in the statement of the Theorem, and the second can be made as small as we wish again using (54).

Thus, provided  $c$  is small enough to effectively use (54), and we are close enough to  $\Psi_{a,b}$  with our initial datum  $\phi$ , estimates (55) and (56) hold, so that

$$d_H(O(\phi), O(\Psi_{a,b})) \leq d_H(O(\phi), O(\psi_{a,b})) + d_H(O(\psi_{a,b}), O(\Psi_{a,b})) < \epsilon.$$

□

## 4 Appendix

### 4.1 Proof of Proposition 3.1

Differently from the mostly used technique of Lagrange multipliers (see e.g. [26]), we here prefer working locally on the constraint, thus we make use of a local parametrization of the manifold with its tangent space. In this way, we still have a functional defined over a linear euclidean space. After this preliminary operation, the problem is treated with the usual IFT (see [1] and [14]). The geometric part of the proof is trivial since the phase space is of finite dimension  $2N$ . However, the estimates are uniform with  $N$ .

We consider the tangent space in a point  $\tilde{z} \in \mathcal{S}$ , as defined by  $T_{\tilde{z}}\mathcal{S} := \{Y \in \mathcal{P} \mid \sum_j \tilde{z}_j Y_j = 0\} = \langle \tilde{z} \rangle^\perp$ , where  $\langle \tilde{z} \rangle$  represents the linear space generated by  $\tilde{z}$ . Since  $X_{H_\Omega}(\tilde{z}) \in T_{\tilde{z}}\mathcal{S}$ , the set  $V := \{Y \in T_{\tilde{z}}\mathcal{S} \mid \sum_j Y_j(X_{H_\Omega}(\tilde{z}))_j = 0\} \subset T_{\tilde{z}}\mathcal{S}$  is a linear subspace of  $\mathcal{P}$  (of dimension  $2N - 2$ ).

<sup>11</sup>Unless one decides to take  $\epsilon \ll R^2(Rr)^r$ , which is not necessary to get a meaningful result of orbital stability. In that case, one would get the stability only of a piece of the periodic orbit.



Take  $\tilde{z} = \tilde{\psi}_0$  as in (39). The phase space  $\mathcal{P}$  can be decomposed into the direct sum of the tangent space  $T_{\tilde{\psi}_0}\mathcal{S}$  and its orthogonal direction  $\tilde{\psi}_0$ , and also the tangent space itself can be decomposed into the field direction  $X_{H_\Omega}(\tilde{\psi}_0)$  and its orthogonal complement  $V$

$$\mathcal{P} = T_{\tilde{\psi}_0}\mathcal{S} \oplus \tilde{\psi}_0, \quad T_{\tilde{\psi}_0}\mathcal{S} = V \oplus X_{H_\Omega}(\tilde{\psi}_0).$$

This gives the characterization

$$V = \{(\tilde{x}, \tilde{y}) \in \mathcal{P} \mid \tilde{x}_0 = \tilde{y}_0 = 0\}. \quad (57)$$

Let us work locally on a neighborhood of  $\tilde{\psi}_0 \in \mathcal{S}$ . There exist  $\mathcal{U}(\tilde{\psi}_0) \subset \mathcal{S}$  and a function  $f : \mathcal{W} \subset T_{\tilde{\psi}_0}\mathcal{S} \rightarrow \langle \tilde{\psi}_0 \rangle$  such that, for any  $\tilde{z} \in \mathcal{U}$  there exists  $h \in T_{\tilde{\psi}_0}\mathcal{S}$  satisfying

$$\tilde{z} = P(h) := \tilde{\psi}_0 + h + f(h);$$

in rough words, locally the sphere is the graph of a function  $f$  defined on the tangent space. The above map is a  $\mathcal{C}^2(\mathcal{W}, \mathcal{U})$  diffeomorphism. From the previous decomposition of  $T_{\tilde{\psi}_0}\mathcal{S}$ , it is locally well defined the submanifold

$$\mathcal{M} := \{\tilde{z} \in \mathcal{U} \mid \tilde{z} = P(h), h \in V \cap \mathcal{W}\}. \quad (58)$$

By construction we have  $T_{\tilde{\psi}_0}\mathcal{M} = V$ .

Since  $H_\Omega$  is a preserved quantity for  $\mathcal{Z}$ , the flow of  $X_{H_\Omega}$  is a symmetry and then

$$d\mathcal{Z}|_{\mathcal{S}}(\tilde{z}, a, b) = 0 \quad \Leftrightarrow \quad d\mathcal{Z}|_{\mathcal{M}}(\tilde{z}, a, b) = 0. \quad (59)$$

We are interested in the problem

$$d\mathcal{Z}|_{\mathcal{M}}(\tilde{z}, a, b) = 0,$$

which has the solution  $\tilde{\psi}_0$  for  $a = b = \mu = 0$ . From the local linear representation of  $\mathcal{M}$ , we can consider  $\mathcal{Z}$  on the linear space  $V$

$$\mathcal{Z}(h, a, b) := \mathcal{Z}(P(h), a, b) = \mathcal{Z}|_{\mathcal{M}}(\tilde{z}, a, b) \quad h \in V \cap \mathcal{W},$$

which is at least  $\mathcal{C}^2(V, \mathbb{R})$ . We look for a map

$$g : (a, b) \in [0, c^*) \times [0, c^*) \mapsto h = g(a, b) \in V, \quad g(0, 0) = 0, \quad (60)$$

such that  $\mathcal{Z}'_h(g(a, b), a, b) = 0$ ; we already know that  $\mathcal{Z}'_h(0, 0, 0) = 0$ .

We set the operator  $F$

$$F(h, a, b) := \mathcal{Z}'_h(h, a, b) = \mathcal{Z}'_\psi(P(h), a, b)P'(h), \quad (61)$$

which, due to Lemma 2.1, is  $\mathcal{C}^1$  from  $V \times \mathbb{R}^2$  to  $V^* := L(V, \mathbb{R})$  and satisfies  $F(0, 0, 0) = \mathcal{Z}'_\psi(\tilde{\psi}_0, 0, 0)\mathbb{I} = 0$ .

The differential  $F'_h(0, 0, 0) = \mathcal{Z}''_\psi(\tilde{\psi}_0, 0, 0)\mathbb{I}$ , which maps  $V$  to its dual  $V^*$ , has the inverse bounded by the constant  $1/C_{h_1}\rho^2$ ; indeed, we already observed in subsection 3.1.1 that, when  $a = b = 0$ , the whole orbit generated by  $\tilde{\psi}_0$  is a constrained strong extremizer, hence the constrained Hessian is coercive in all the directions transverse to  $X_{H_\Omega}(\tilde{\psi}_0)$ , with coercivity constant  $2C_{h_1}\rho^2$  in the euclidean norm. A direct computation, which is based on the explicit computations developed in subsection 3.1.1, shows indeed that one has

$$|\mathcal{Z}''_\psi(\tilde{\psi}_0, 0, 0)\mathbb{I}[Y, Y]| \geq C_2 \sum_{j \neq 0} Y_j^2 \geq C_2 \|Y\|^2, \quad C_2 := 2C_{h_1}\rho^2, \quad (62)$$

hence

$$\left\| [F'_h(0, 0, 0)]^{-1} \right\|_{L(V^*, V)} \leq 1/C_2. \quad (63)$$

Then there exist  $\mu_2^*(\rho)$ , and hence from (21) a  $c_2^* = c_2^*(\rho)$ , and a function  $g \in \mathcal{C}^1([0, c_2^*] \times [0, c_2^*], V)$  as in (60) such that  $F(g(a, b), a, b) = 0$  with

$$\|g(a, b)\| < \mu, \quad |\mu| < \mu_2^*.$$

Furthermore, we recall that the IFT is based on the contraction Theorem on a closed  $\epsilon$ -ball  $B_\epsilon \subset V$  for the operator  $A_{a,b}(h) := h - [F'_h(0, 0, 0)]^{-1}F(h, a, b) : V \mapsto V$ . The requirements of being a contraction and surjective on  $B_\epsilon$ , implies that  $\mu_2^*$  is bounded by the coercive constant  $C_2$  in (63).

From the property  $f(h) = o(\|h\|)$  of the parametrization  $P$ , it immediately follows that the solution  $\tilde{\psi}_{a,b} := P(g(a, b)) =: G(a, b)$  is  $\mu$  close to  $\tilde{\psi}_0$

$$\|\tilde{\psi}_{a,b} - \tilde{\psi}_0\| \leq C\mu.$$

□

## 4.2 Proof of Proposition 3.2

From the continuity of  $\mathcal{Z}''$  we deduce that  $\tilde{\psi}_{a,b}$  is still a strong extremizer in the direction  $V$ , with a coercive constant  $C_\mu = \mathcal{O}(\rho^2)$ . Hence the orbit generated by the flow of  $X_{H_\Omega}(\psi)$  is orbitally Lyapunov stable with  $\mathcal{Z}$  being the Lyapunov function (see [3, 5, 25, 26]).

In few words (inspired also by Section 8 of [5], in particular Lemmas 8.5 and 8.6, although we work in the simplified case of a finite dimensional phase space), given a generic point of the orbit  $\tilde{\eta} \in \mathcal{O}(\tilde{\psi}_{a,b})$ , there exists a neighborhood  $\mathcal{W}_0$  of  $\tilde{\eta}$  where a suitable set of coordinates can be introduced. This local representation is based on the decomposition  $P_{\tilde{\eta}} = \nabla H_\Omega(\tilde{\eta}) \oplus V_{\tilde{\eta}}$  of the hyperplane  $P_{\tilde{\eta}}$  orthogonal to  $X_{H_\Omega}(\tilde{\eta})$ , for any  $\tilde{\eta} \in \mathcal{O}(\tilde{\psi}_{a,b})$ . More precisely, there exists a (tubular) neighborhood  $\mathcal{W}_0$  of  $\tilde{\eta}$  such that, for any point  $\tilde{z} \in \mathcal{W}_0$ , the hyperplane through  $\tilde{z}$  and orthogonal to  $\mathcal{O}(\tilde{\psi}_{a,b})$  is unique. This plane intersects the periodic orbit  $\mathcal{O}(\tilde{\psi}_{a,b})$  in a point  $\tilde{\xi}$ , which can be obtained as the evolution of  $\tilde{\eta}$  at “time”  $\varphi$  along the flow of the periodic orbit. Hence, using the previous notation, such a plane can be decomposed as  $P_{\tilde{\xi}} = \nabla H_\Omega(\tilde{\xi}) \oplus V_{\tilde{\xi}}$ . This implies that  $\tilde{z}$  can be locally represented by the coordinates

$$\tilde{z} \equiv (\varphi, E, v) \in \mathbb{R} \times \mathbb{R} \times V_{\tilde{\xi}}, \quad (64)$$

where  $E$  represents the displacement in the  $\nabla H_\Omega(\tilde{\xi})$  direction and  $v$  the displacement in the  $V_{\tilde{\xi}}$  direction(s). Using these local coordinates in order to represent  $\tilde{z} = \tilde{\phi}(t) \in \mathcal{W}_0$ , the orbital distance of  $\tilde{\phi}(t)$  from  $\mathcal{O}(\tilde{\psi}_{a,b})$  is controlled in  $\mathcal{W}_0$  by

$$d(\tilde{\phi}(t), \mathcal{O}(\tilde{\psi}_{a,b})) \leq \inf_{w \in \mathcal{O}(\tilde{\psi}_{a,b}) \cap \mathcal{W}_0} \|w - \tilde{\phi}(t)\| \leq c_1 |E(t)| + c_2 \|v(t)\|, \quad (65)$$

with  $c_{1,2}$  depending on  $\mathcal{W}_0$ . The first term  $|E(t)|$  represents the variation of  $|H_\Omega(\tilde{\phi}(t)) - H_\Omega(\tilde{\psi}_{a,b})| = |H_\Omega(\tilde{\phi}(t)) - H_\Omega(\tilde{\xi})|$ : indeed, being  $E(t)$  the coordinate associated to the direction  $\nabla H_\Omega(\tilde{\xi})$ , with  $\tilde{\xi} \in \mathcal{O}(\tilde{\psi}_{a,b})$ , it controls the displacement orthogonal to  $\mathcal{S}_{\tilde{\xi}}$ . The second term  $\|v(t)\|$  is instead related to the variation of  $|\mathcal{Z}(\tilde{\phi}(t)) - \mathcal{Z}(\tilde{\psi}_{a,b})| = |\mathcal{Z}(\tilde{\phi}(t)) - \mathcal{Z}(\tilde{\xi})|$ , which controls the  $V_{\tilde{\xi}}$  directions transverse to the orbit, *provided*  $\|v(t)\|$  is *small enough*. Here enters the fact that any point  $\tilde{\xi} \in \mathcal{O}(\tilde{\psi}_{a,b})$  is a local extremizer for  $\mathcal{Z}$  constrained to<sup>12</sup>  $\mathcal{M}_{\tilde{\xi}}$ . Indeed, if we take a point  $\tilde{z} \in V_{\tilde{\xi}}$  close enough to  $\tilde{\xi}$  (such that  $\mathcal{Z}$  almost coincides with its quadratic part), then a Taylor expansion gives

$$\mathcal{Z}(\tilde{z}) - \mathcal{Z}(\tilde{\xi}) = \frac{1}{2} \mathcal{Z}''(\tilde{\xi})[\tilde{z} - \tilde{\xi}, \tilde{z} - \tilde{\xi}] + h.o.t.,$$

which provides the bound

$$\|\tilde{z} - \tilde{\xi}\|^2 = \|v\|^2 \leq \frac{3}{C_\mu} |\mathcal{Z}(\tilde{z}) - \mathcal{Z}(\tilde{\xi})|, \quad \|v\| \ll C_\mu \sim \rho^2. \quad (66)$$

<sup>12</sup>One can define  $\mathcal{M}_{\tilde{\xi}}$  as the submanifold tangent to  $V_{\tilde{\xi}}$  as in (58).

Thus there exists a neighborhood  $\mathcal{W}_1 \subset \mathcal{W}_0$  of  $\tilde{\eta}$  such that if  $\tilde{\phi}(t) \in \mathcal{W}_1$  then (65) becomes

$$d(\tilde{\phi}(t), \mathcal{O}(\tilde{\psi}_{a,b})) \leq \sqrt{c_3 |H_\Omega(\tilde{\phi}(t)) - H_\Omega(\tilde{\psi}_{a,b})| + \frac{c_4}{C_\mu} |\mathcal{Z}(\tilde{\phi}(t)) - \mathcal{Z}(\tilde{\psi}_{a,b})|}. \quad (67)$$

with  $c_{3,4}$  depending on  $\mathcal{W}_1$ . Since  $\mathcal{O}(\tilde{\psi}_{a,b})$  is compact (being homeomorphic to  $S^1$ ), we can cover a whole neighborhood  $\mathcal{U}$  of this orbit with a finite collection (independent of  $N$ ) of local neighborhoods like  $\mathcal{W}_1$  and set of coordinates like (64), such that (67) holds true. Since both  $H_\Omega$  and  $\mathcal{Z}$  are continuous (analytic, see Lemma 2.1) constants of motion for  $K$ , the requirement of staying in  $\mathcal{U}$  is translated in a closeness condition for the initial datum  $\tilde{\phi}(0)$ : there exists  $\delta(\epsilon)$  such that

$$d(\tilde{\phi}(0), \mathcal{O}(\tilde{\psi}_{a,b})) < \delta \quad \Rightarrow \quad \sqrt{c_3 |H_\Omega(\tilde{\phi}(0)) - H_\Omega(\tilde{\psi}_{a,b})| + \frac{c_4}{C_\mu} |\mathcal{Z}(\tilde{\phi}(0)) - \mathcal{Z}(\tilde{\psi}_{a,b})|} < \epsilon.$$

This actually gives  $d(\mathcal{O}(\tilde{\phi}), \mathcal{O}(\tilde{\psi}_{a,b})) < \epsilon$ , i.e. the orbit we aim to control is contained in the tubular neighborhood of the breather  $\tilde{\psi}_{a,b}$  for the normal form  $K$ . To conclude the proof we also need the symmetric control, to avoid that our orbit, despite being in  $\mathcal{U}$ , does not actually follow the whole trajectory of the breather. Indeed, in full generality it could happen that the orbit goes back and forth only in a section of  $\mathcal{U}$ ; or it could happen that such a neighborhood is not homotopic to an  $S^1$ , e.g. it has an “eight” shape, and in that case the orbit could use the “connection” as a shortcut to follow only a part of the orbit without leaving  $\mathcal{U}$ . In our case these problems do not arise: indeed the GdNLS breather is given by the action of  $e^{i\lambda t}$  on  $\mathcal{S}$ , i.e. it is a maximal circle on a sphere whose radius is of order  $\rho$ . On the other hand,  $\mathcal{U}$  is the cartesian product of the breather and a disc of co-dimension one, whose radius has to be of order  $\epsilon$  which is constrained to be smaller than  $\rho^2$ . As a first consequence  $\mathcal{U}$  is necessarily homotopic to an  $S^1$ . Moreover the component of the vector field transverse to the disc is a small perturbation of the vector field in the point of the breather orbit which lies in the disc itself. It is thus not possible for any orbit in  $\mathcal{U}$  to stop flowing along the tubular neighborhood, and this happen in a time which is a small perturbation of the period of the breather.

The above arguments allow us to get also  $d(\mathcal{O}(\tilde{\psi}_{a,b}), \mathcal{O}(\tilde{\phi})) < \epsilon$ , and this concludes the proof.  $\square$

**Acknowledgments:** We warmly thank Dario Bambusi for suggesting to separate this application from the normal form result [17], and for other useful comments. This research is partially supported by MIUR-PRIN program under project 2010 JJ4KPA (“Teorie geometriche e analitiche dei sistemi Hamiltoniani in dimensioni finite e infinite”).

## References

- [1] A. AMBROSETTI AND G. PRODI, *A primer of nonlinear analysis*, Cambridge University Press, Cambridge, 1995. Corrected reprint of the 1993 original.
- [2] S. AUBRY, *Breathers in nonlinear lattices: existence, linear stability and quantization*, Phys. D, 103 (1997), pp. 201–250. Lattice dynamics (Paris, 1995).
- [3] D. BAMBUSI, *Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators*, Nonlinearity, 9 (1996), pp. 433–457.
- [4] ———, *On long time stability in Hamiltonian perturbations of non-resonant linear PDEs*, Nonlinearity, 12 (1999), pp. 823–850.
- [5] D. BAMBUSI AND N. N. NEKHOROSHEV, *A property of exponential stability in nonlinear wave equations near the fundamental linear mode*, Phys. D, 122 (1998), pp. 73–104.

- [6] R. CARRETERO-GONZÁLEZ, J. D. TALLEY, C. CHONG, AND B. A. MALOMED, *Multistable solitons in the cubic-quintic discrete nonlinear Schrödinger equation*, Phys. D, 216 (2006), pp. 77–89.
- [7] C. CHONG, R. CARRETERO-GONZÁLEZ, B. A. MALOMED, AND P. G. KEVREKIDIS, *Variational approximations in discrete nonlinear Schrödinger equations with next-nearest-neighbor couplings*, Phys. D, 240 (2011), pp. 1205–1212.
- [8] C. CHONG AND D. E. PELINOVSKY, *Variational approximations of bifurcations of asymmetric solitons in cubic-quintic nonlinear Schrödinger lattices*, Discrete Contin. Dyn. Syst. Ser. S, 4 (2011), pp. 1019–1031.
- [9] T. DAUXOIS, M. PEYRARD, AND C. R. WILLIS, *Localized breather-like solution in a discrete Klein-Gordon model and application to DNA*, Phys. D, 57 (1992), pp. 267–282.
- [10] A. GIORGILLI, S. PALEARI, AND T. PENATI, *Extensive adiabatic invariants for nonlinear chains*, J. Statist. Phys., 148 (2012), pp. 1106–1134.
- [11] ———, *An extensive adiabatic invariant for the Klein-Gordon model in the thermodynamic limit*, Ann. Henri Poincaré, online first (2014).
- [12] G. JAMES, P. G. KEVREKIDIS, AND J. CUEVAS, *Breathers in oscillator chains with hertzian interactions*, Physica D: Nonlinear Phenomena, 251 (2013), pp. 39 – 59.
- [13] N. KARACHALIOS, B. SÁNCHEZ-REY, P. KEVREKIDIS, AND J. CUEVAS, *Breathers for the discrete nonlinear schrödinger equation with nonlinear hopping*, Journal of Nonlinear Science, 23 (2013), pp. 205–239.
- [14] A. N. KOLMOGOROV AND S. V. FOMIN, *Elementy teorii funktsii i funktsionalnogo analiza*, “Nauka”, Moscow, sixth ed., 1989. With a supplement, “Banach algebras”, by V. M. Tikhomirov.
- [15] V. KOUKOULOYANNIS, P. G. KEVREKIDIS, J. CUEVAS, AND V. ROTHOS, *Multibreathers in Klein-Gordon chains with interactions beyond nearest neighbors*, Phys. D, 242 (2013), pp. 16 – 29.
- [16] R. S. MACKAY AND S. AUBRY, *Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators*, Nonlinearity, 7 (1994), pp. 1623–1643.
- [17] S. PALEARI AND T. PENATI, *An extensive resonant normal form for an arbitrarily large Klein-Gordon model*, Ann. Mat. Pura Appl. (4), (to appear). <http://arxiv.org/abs/1404.2730>.
- [18] D. PELINOVSKY AND A. SAKOVICH, *Multi-site breathers in Klein-Gordon lattices: stability, resonances and bifurcations*, Nonlinearity, 25 (2012), p. 3423.
- [19] D. E. PELINOVSKY, P. G. KEVREKIDIS, AND D. J. FRANTZESKAKIS, *Stability of discrete solitons in nonlinear Schrödinger lattices*, Phys. D, 212 (2005), pp. 1–19.
- [20] T. PENATI AND S. PALEARI, *Breathers and Q-breathers: two sides of the same coin*, SIAM J. Appl. Dyn. Syst., 11 (2012), pp. 1–30.
- [21] M. PEYRARD AND A. R. BISHOP, *Statistical mechanics of a nonlinear model for dna denaturation*, Phys. Rev. Lett., 62 (1989), pp. 2755–2758.
- [22] W.-X. QIN AND X. XIAO, *Homoclinic orbits and localized solutions in nonlinear Schrödinger lattices*, Nonlinearity, 20 (2007), pp. 2305–2317.
- [23] P. ROSENAU AND S. SCHOCHET, *Compact and almost compact breathers: a bridge between an anharmonic lattice and its continuum limit*, Chaos, 15 (2005), pp. 015111, 18.

- [24] P. TCHOFO DINDA AND M. REMOISSENET, *Breather compactons in nonlinear klein-gordon systems*, Phys. Rev. E, 60 (1999), pp. 6218–6221.
- [25] M. I. WEINSTEIN, *Lyapunov stability of ground states of nonlinear dispersive evolution equations*, Comm. Pure Appl. Math., 39 (1986), pp. 51–67.
- [26] ———, *Excitation thresholds for nonlinear localized modes on lattices*, Nonlinearity, 12 (1999), pp. 673–691.
- [27] K. YOSHIMURA, *Stability of discrete breathers in nonlinear Klein-Gordon type lattices with pure anharmonic couplings*, J. Math. Phys., 53 (2012), pp. 102701, 20.